



Prediction of the Shear Wave Velocity from Compressional Wave Velocity for Gachsaran Formation

Saeed PARVIZI¹, Riyaz KHARRAT¹, Mohammad R. ASEF²,
Bijan JAHANGIRY³, and Abdolnabi HASHEMI¹

¹Petroleum University of Technology, Department of Petroleum Engineering,
Ahwaz, Iran; e-mail: saeed.parvizi68@gmail.com

²Kharazmi University, Faculty of Earth Sciences, Department of Applied Geology,
Tehran, Iran

³National Iranian South Oil Company, Department of Drilling Engineering,
Ahwaz, Iran

A b s t r a c t

Shear and compressional wave velocities, coupled with other petrophysical data, are very important for hydrocarbon reservoir characterization. *In situ* shear wave velocity (V_s) is measured by some sonic logging tools. Shear velocity coupled with compressional velocity is vitally important in determining geomechanical parameters, identifying the lithology, mud weight design, hydraulic fracturing, geophysical studies such as VSP, *etc.* In this paper, a correlation between compressional and shear wave velocity is obtained for Gachsaran formation in Maroon oil field. Real data were used to examine the accuracy of the prediction equation. Moreover, the genetic algorithm was used to obtain the optimal value for constants of the suggested equation. Furthermore, artificial neural network was used to inspect the reliability of this method. These investigations verify the notion that the suggested equation could be considered as an efficient, fast, and cost-effective method for predicting V_s from V_p .

Key words: compressional wave velocity, shear wave velocity, sonic log, DSI log, MATLAB software.

1. INTRODUCTION

Compressional and shear velocities (V_p and V_s , respectively) are important in seismic inversion and petrophysical evaluation of formations, especially for analysis of reservoir geomechanical properties. For estimating the geomechanical parameters such as Young's modulus, Poisson's ratio, and Lamé parameters, both V_p and V_s plus density are needed. Accordingly, knowing V_p , V_s , and density, other elastic parameters of a rock formation can be calculated in terms of the acoustic wave velocities (Liu *et al.* 2012).

$$\mu = \rho V_s^2, \quad (1)$$

$$E = \frac{\rho V_s^2 (3V_p^2 - 4V_s^2)}{(V_p^2 - V_s^2)}, \quad (2)$$

$$\lambda = \rho V_p^2 - 2\rho V_s^2, \quad (3)$$

$$K = \rho V_p^2 - \frac{4}{3}\rho V_s^2, \quad (4)$$

$$\nu = \frac{(V_p^2 - 2V_s^2)}{2(V_p^2 - V_s^2)}. \quad (5)$$

According to the acoustic wave propagation theory, the P -wave (V_p) and S -wave (V_s) velocities can be expressed as:

$$V_p = \sqrt{\frac{\lambda + 2\mu}{\rho}}, \quad (6)$$

$$V_s = \sqrt{\frac{\mu}{\rho}}. \quad (7)$$

The above-mentioned geomechanical parameters are useful in estimating maximum and minimum horizontal stresses, mud weight design, *etc.* Therefore, rock mechanical properties can be estimated using some sonic log providing P - and S -wave velocity information such as dipole sonic log and so on. However, very often S -wave velocity is not recorded in the field all the time due to the cost constraints and lack of technology. Therefore, prediction of the S -wave velocity is an interesting objective for researchers (Farrokhrouz and Asef 2010). Alternatively, if sonic tools to measure V_s are not available, we may use a prediction equation for estimating shear wave velocity based on compressional wave velocity obtained from monopole sonic log (Liu *et al.* 2012). Almost all such equations are empirical (Asef

and Farrokhrouz 2010). Castagna *et al.* (1985) suggested Eq. 8, that is, a linear equation based on laboratory data for water-saturated elastic silicate rocks.

$$V_s = 0.862 V_p - 1.172 . \quad (8)$$

Han (1986) used an extensive experimental dataset of sandstone with wide ranges of porosity and clay content variation to obtain the equation

$$V_s = 0.794 V_p - 0.787 . \quad (9)$$

Brocher (2005) developed a relation between elastic wave velocity in the Earth's crust and developed Eq. 10 indicating a nonlinear relationship between V_p and V_s .

$$V_s = 0.7858 - 1.2344 V_p + 0.7949 V_p^2 - 0.1238 V_p^3 + 0.0064 V_p^4 , \quad 1.5 < V_p < 8 , \quad (10)$$

where V_s and V_p are in km/s. It is noticed that each of the above empirical equations was developed for a specific field and a specific lithology. Nevertheless, if they are used at Iranian fields, they may result in erroneous predictions. Figures 1 to 3 illustrate the accuracy of the said correlations in Gachsaran formation.

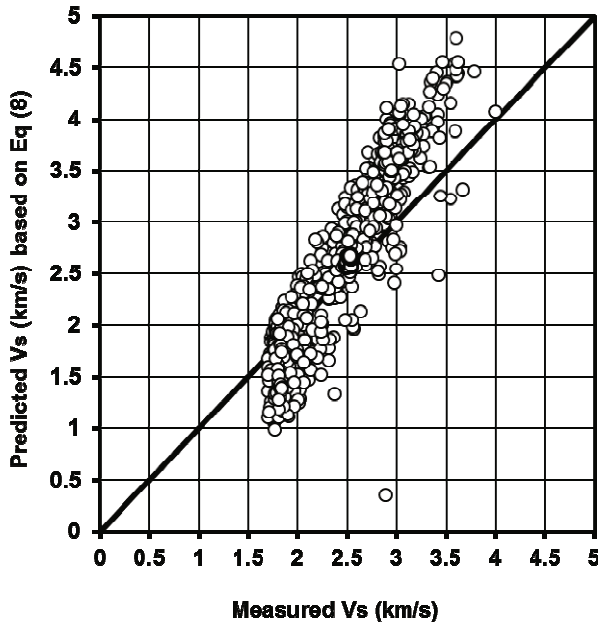


Fig. 1. Calculated V_s versus measured values using Eq. 8.

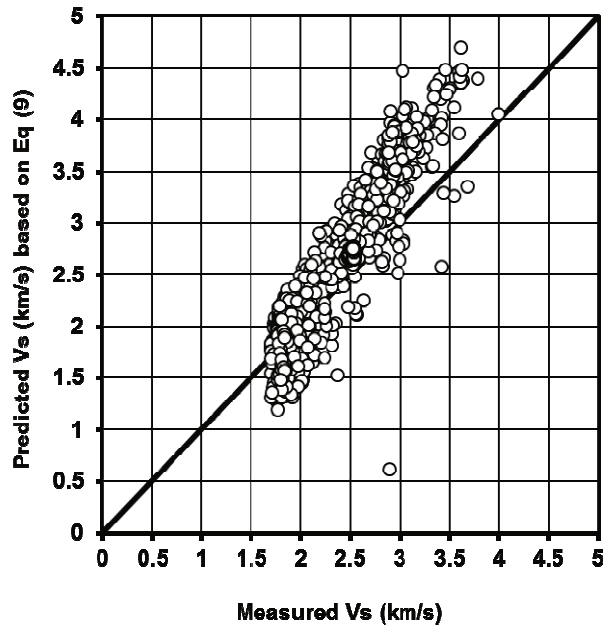


Fig. 2. Calculated V_s versus measured values using Eq. 9.

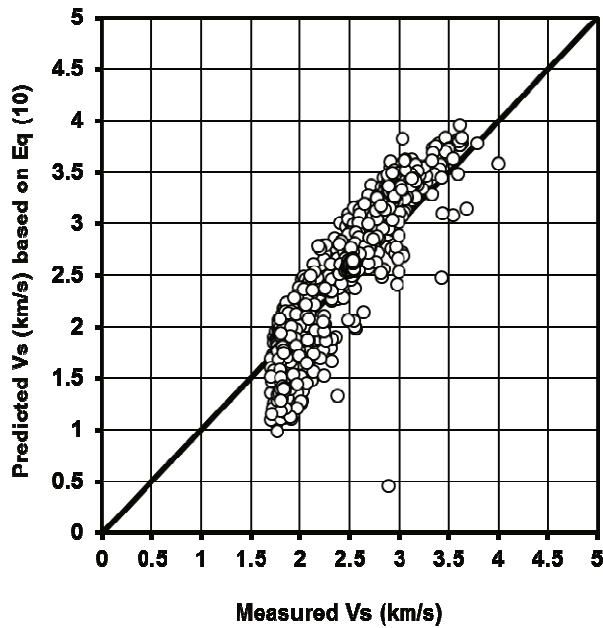


Fig. 3. Calculated V_s versus measured values using Eq. 10.

Therefore, the present study was conducted to develop a reliable model for predicting shear wave velocity in a rock, based on compressional wave velocity in Maroon oil field. For this purpose, we utilized different regression models to obtain the most appropriate approach. We compared real field measurements of V_s with predicted values. Furthermore, we applied a neural network model to examine the accuracy and reliability of the suggested approach.

2. GEOLOGICAL SETTINGS OF THE STUDIED FIELDS

In this research we used data from Maroon oil field in Ahwaz, Southeast Iran. The main objective of this study was predicting shear wave velocity based on P -wave data. As it is obvious from petrophysical logs and geological studies, Gachsaran formation in Maroon oil field consist of seven members as presented in Table 1 (Memari 2013).

Table 1

Lithology of the studied formations

Formation	Lithology	Density
Gachsaran 7	made up of mainly anhydrite and some grey marl and limestone	$2.58 < \text{density} < 3.02$
Gachsaran 6	mainly anhydrite, salt, red and gray marl layers	$2.46 < \text{density} < 2.9$
Gachsaran 5	mainly anhydrite, salt, red and gray marl layers	$2.22 < \text{density} < 2.85$
Gachsaran 4	mainly anhydrite, salt and gray marl layers	$2.5 < \text{density} < 2.97$
Gachsaran 3	thick anhydrites with subordinate salt in the lower half, and alternating anhydrites, thin limestones and marls in the upper half	$2.46 < \text{density} < 2.94$
Gachsaran 2	thick salt units with intervening anhydrite and thin limestones	$2.53 < \text{density} < 2.92$
Gachsaran 1 (Cap Rock)	mainly anhydrite and gray marl and minor layers of limestone	$2.69 < \text{density} < 2.71$

The dipole shear sonic imager (DSI) combines monopole and dipole sonic acquisition capabilities for the reliable acoustic measurement of compressional, shear and Stoneley slowness. This is a Schlumberger company tool. DSI log for Gachsaran formation in Maroon field was run. Therefore, V_s and V_p were obtained directly from this log.

3. DATA ANALYSIS

Regression analysis was carried out to estimate and model the relationship between a response variable and one or more predictors. An empirical equation is not a perfect relationship. In general, all observation points for a statistical relationship do not fall directly on the prediction curve. Meanwhile, there are many regression models to define the best fit between two parameters, such as: linear, polynomial, exponential, *etc.* (Rawlings *et al.* 1998). Therefore, in the first step we obtained well log readings of V_s and V_p for 4000 points from 2984 to 3617 m depth at the studied well. Randomly, we picked 1/4 of data points (1000 points) as “additional validating data” for further investigation, and we used 3/4 of data points (3000 points) as “modeling data” for developing prediction equation.

TableCurve software package and MATLAB Curve Fitting Toolbox were used to carry out all regression analyses (linear, polynomial, exponential...). Therefore, the mutual relationship between V_p and V_s was studied. The value of R^2 and adjusted R^2 (R^2 adjusted for the number of independent variables in the model) are presented in Table 2. R^2 is the *r*-square defined as:

$$R^2 = 1 - \frac{\sum_i (y_i - f_i)^2}{\sum_i (y_i - \bar{y})^2} \quad (11)$$

A data set has values y_i which have an associated modeled value f_i (sometimes called the predicted values) and \bar{y} is the mean of the observed data. A very good fit yields value of 1, whereas a poor fit results in a value near 0. Adjusted R^2 is defined as:

$$R^2_{\text{adj}} = 1 - \frac{\frac{\sum_i (y_i - f_i)^2}{n-d-1}}{\frac{\sum_i (y_i - \bar{y})^2}{n-1}} \quad (12)$$

where n is the number of points in the data set. Adjusted R^2 is especially important in this research because if unnecessary variables are included, R^2 can be misleadingly high. As it is observed in Table 2, in case of Eq. 19 the value of R^2 increased and RMSE reduced significantly. Also, it should be clarified that R^2 and adjusted R^2 are in reasonable agreement with each other in case of Eq. 19 which implies contribution of V_p to the prediction equation. Figure 4 illustrates the scatter of data points against prediction curve for Eq. 19. The accuracy of Eq. 19 is further clarified when measured values are plotted against predicted values in Fig. 5. It seems that the exponential model (Eq. 19) is the best for Gachsaran formation. The results of these

analyses are listed in Table 2. The Root Mean Squared Error (RMSE) in all models was estimated as:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (V_s(\text{observed})_i - V_s(\text{predicted})_i)^2}{n}} \quad (13)$$

Table 2
Prediction equations and impact parameters for Gachsaran formation

Equation name	Model summary			Equation	
	R^2	R^2_{adj}	RMSE	Statistical criteria	
Linear	0.8885	0.8885	0.16	$V_s = 0.4078 - 0.4614 V_p$	(14)
4th order	0.8835	0.8832	0.1481	$V_s = 14.38 - 12.04 V_p + 4.061 V_p^2 - 0.5692 V_p^3 + 0.02913 V_p^4$	(15)
Power	0.8868	0.8868	0.1612	$V_s = 0.7144 V_p^{0.8308}$	(16)
Logarithmic	0.9014	0.9013	0.1506	$V_s = 15.64 - 36.72 \ln(V_p) + 34.2 (\ln(V_p))^2 + 13.27 (\ln(V_p))^3 - 1.926 (\ln(V_p))^4$	(17)
Inverse	0.8187	0.8186	0.2041	$V_s = 4.226 - \frac{7.465}{V_p}, \quad 1.767 \leq V_p$	(18)
Exponential	0.8946	0.8945	0.1555	$V_s = 1.532 + \frac{1.8790389}{1 + e^{\left(\frac{V_p - 4.4767345}{0.839419}\right)}}$	(19)
S-curve	0.863	0.8629	0.1774	$V_s = e^{\left(1.7 - \frac{3.454}{V_p}\right)}$	(20)

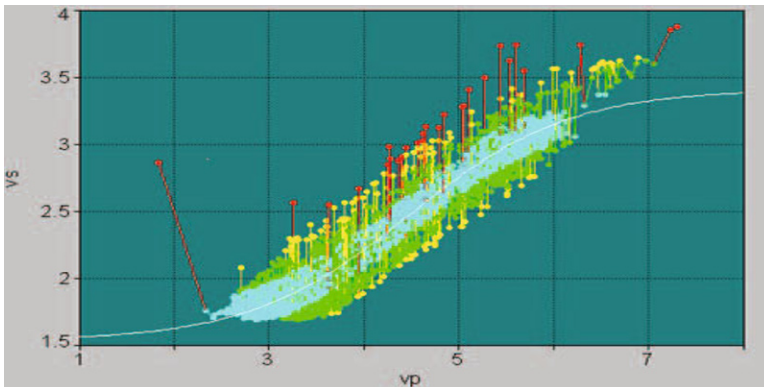


Fig. 4. The main plot of Eq. 19.

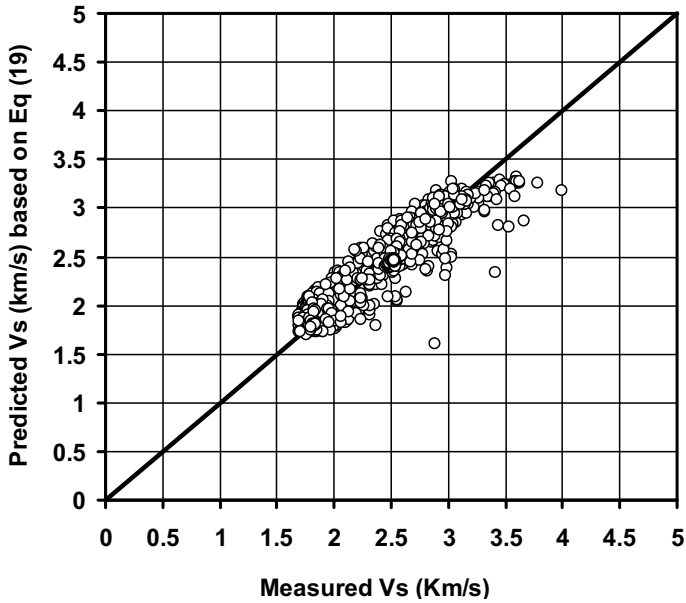


Fig. 5. Measure V_s versus values calculated based on Eq. 19.

4. GENETIC ALGORITHM OPTIMIZATION

Genetic Algorithm (GA) is mostly used for solving optimization problems based on principles of evolution. GA is essentially an iterative, population based, parallel global search algorithm. Solution to the problem is shown by each individual in the population and the population of individuals is kept at each generation. A fitness value is given to any individual to clue the search. Individuals with superior fitness values are selected and undergo genetic transformation by genetic operators such as crossover and mutation. The crossover operator randomly selects two individuals as parents and exchanges part of their structure to produce two new individuals. The mutation operator just randomly selects one individual from the parent population and changes its internal demonstration and puts it in the child population. For improving the search performance, both of the operator rates should be intently adjusted. The newly produced child population becomes the parent population for the subsequent generation and undergoes the same process until a stopping criterion has been satisfied.

For the present study, MATLAB Genetic Algorithm Toolbox for optimization was used to search optimal values of correlation constants for Eq. 19. The performance of GA is generally affected by the size of the initial population, the number of generations, and GA parameters including selection rate, mutation rate, and crossover rate. Therefore, the best fitness of the end

population was evaluated under different sizes of initial population and generations as well as different selection rate, mutation rate, and crossover rate. The best fitness was obtained when the initial population of 2000, number of generations of 200, selection rate of 0.08, mutation rate of 0.05, and crossover rate of 0.6 were obtained by trial and error. The optimum constants are shown in Eq. 21.

$$V_s = 1.68 + \frac{1.56}{1 + \text{EXP}\left(\frac{-V_p - 4.46}{0.63}\right)} \quad (21)$$

5. VERIFICATION ANALYSIS

At this step, V_s is predicted using author's equation (Eq. 21) as well as Eqs. 8-10 without any change of constants in these three equations. RMSE is calculated for identifying the accuracy of each equation and comparing the results of correlations with real value. The results are shown in Table 3. In this table it is clearly observed that RMSE for Eq. 21 is reasonably less than that for other equations.

Table 3
RMSE for different prediction equations

Correlation name	Equation no.	Calculated RMSE
Authors	21	0.1741655
Castagna <i>et al.</i> (1985)	8	0.469219
Han (1986)	9	0.458684
Brocher (2005)	10	0.336383

6. NEURAL NETWORK MODEL

MATLAB Neural Network Toolbox was used to examine the accuracy of Eq. 21. In the first step, a network including 3000 points (out of 4000) was generated; we called it the "modeling data", while 1000 points were the unseen data. We did not use 1000 points to generate neural network model to avoid overtraining. Instead, we used these 1000 points as "additional validating data". We considered V_p data as input and V_s data as output. This is because the objective of this research was predicting V_s based on V_p . Accordingly, "modeling data" (containing 3000 points) were divided into 3 parts: 70% for training (2100 points), 15% for validation (450 points), and 15% for testing (450 points). We know that the number of the hidden neurons affects the accuracy of the result. In order to obtain the best results, we

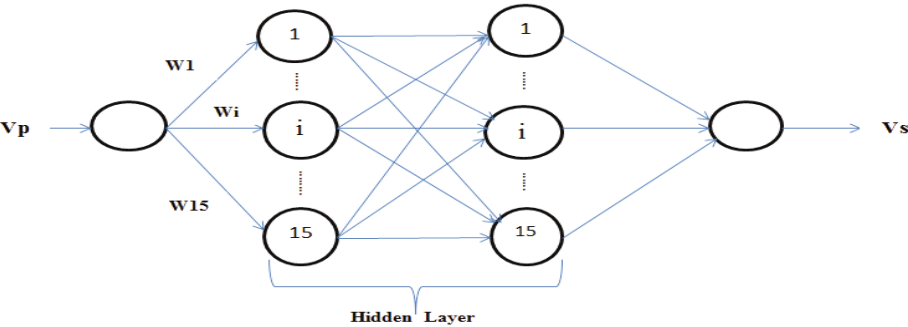


Fig. 6. The architecture of the network developed for V_s prediction.

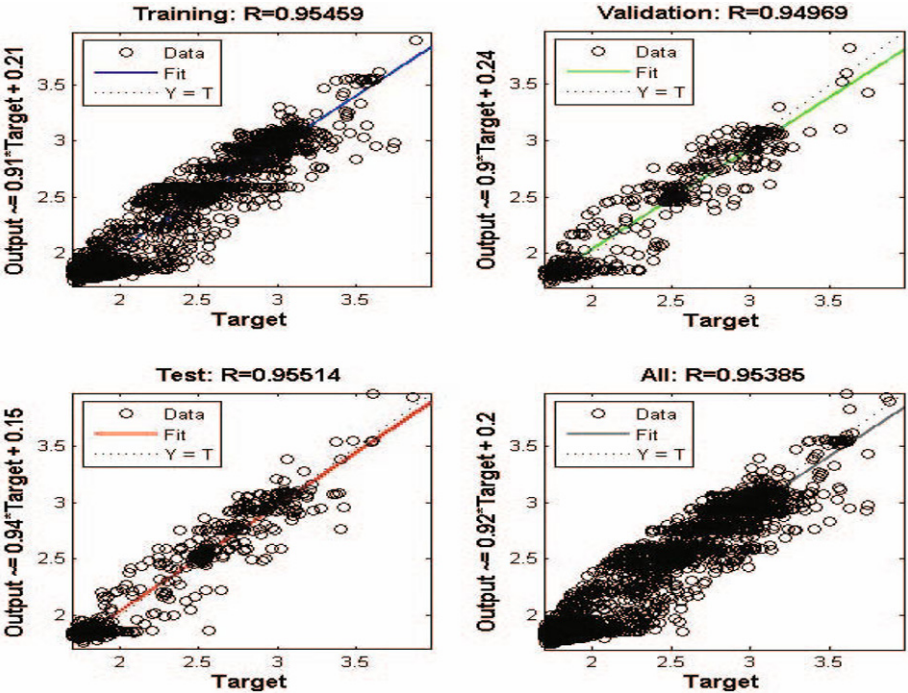


Fig. 7. Establishment of ANN for predicting V_s using neural network fitting.

examined different number of hidden neurons. Finally, the best results were obtained by 30 hidden neurons (two hidden layers with 15 hidden unite in each). Figure 6 represents a simple drawing of the structure of the network made to predict V_s .

Figure 7 illustrates four steps of the network generation and their fitness values. This figure shows that the correlation coefficient for tested data is 0.95514, which indicates that reliable results can be obtained.

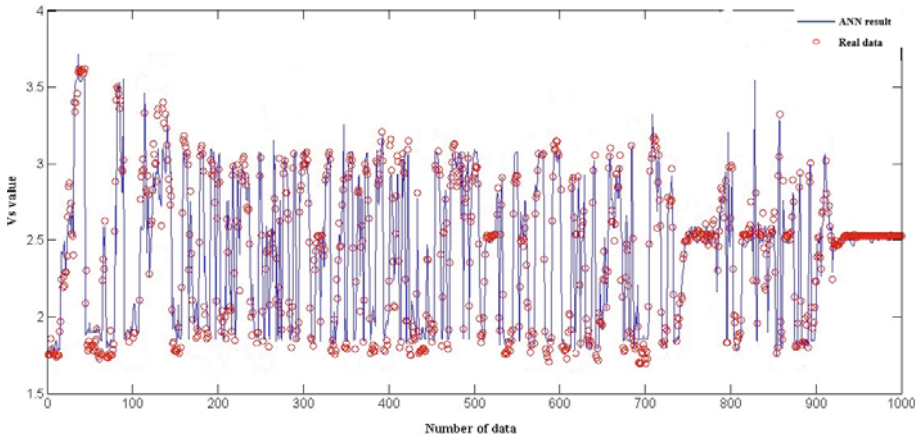


Fig. 8. Performance of ANN for predicting V_s using “additional validating data”.

In the next step, the established network model was employed for predicting V_s using “additional validating data” (1000 points of unseen data). Then the results of the network was compared with real value of the V_s . Figure 8 illustrates the comparison of the V_s value predicted by Artificial Neural Network (ANN) against real values. RMSE and R^2 were equal to 0.1573 and 0.8985, respectively, which indicates a fairly good prediction.

7. DETERMINING THE ACCURACY OF EQUATION 21 IN OTHER WELLS

In the last step of this research, reliability of Eq. 21 was further examined in other wells. Well B at Maroon reservoir has monopole sonic log in Gachsaran formation and only V_p can be obtained from this log. Both Eq. 21 and

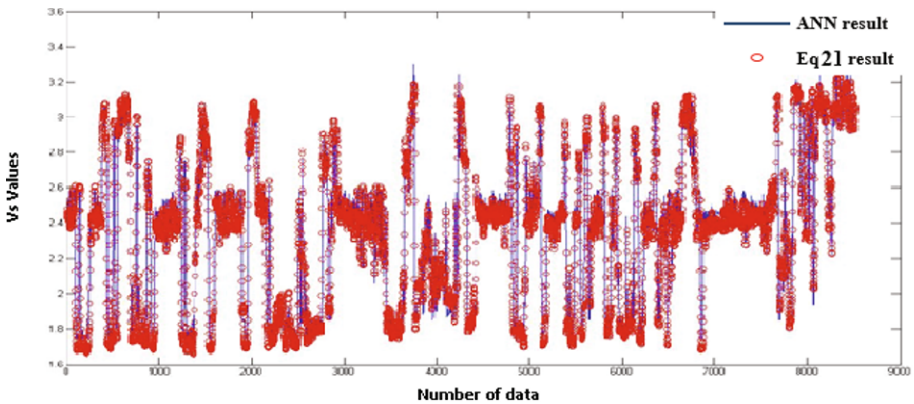


Fig. 9. Prediction of V_s from V_p using ANN/Eq. 21 in well B in Maroon reservoir.

the established ANN were used to predict shear wave velocity for Gachsaran formation in well B. A statistical correlation assessment was carried out to compare predicted shear wave velocity based on these two approaches (Eq. 21 and ANN). RMSE and R^2 were equal to 0.464 and 0.987, respectively. This indicated that predictions of V_s from V_p based on these two methodologies are very similar. The result is shown in Fig. 9. However, Eq. 21 is preferred because it is simpler.

8. CONCLUSIONS

We demonstrated that the shear wave velocity can be predicted from the compressional wave velocity. We suggested an equation that can efficiently minimize the errors and is more accurate than the three previous equations, namely Castagna *et al.* (1985), Han (1986), and Brocher (2005) ones. The suggested equation is reliable to predict shear wave velocity for anhydrite, salt, marl, and limestone formations. Genetic algorithm was used to obtain the optimal value for constants of the suggested equation. Accordingly, both the ANN model and the suggested equation could successfully predict V_s . The suggested correlations for predicting the shear velocity in Gachsaran formation is reliable enough to be used in a case monopole log is at hand and V_S data are not available due to the cost constrains, lack of technology or old cased wells.

Acknowledgement. Authors would like to appreciate the National Iranian Oil Company (NIOC) for providing data and support.

Nomenclature

DSI Log	– Dipole Shear Sonic imager
MDA	– Monopole-Dipole Array
VSP	– Vertical Seismic Profiling
V_s	– Shear wave velocity [km/s]
V_p	– Compressional wave velocity [km/s]
E	– Young's modulus
ν	– Poisson ratio
k	– Bulk modulus
λ	– Lamé parameter
μ	– Shear modulus
R^2	– Square error
R^2_{adj}	– Adjusted Square Error
RMSE	– Root Mean Square Error
ANN	– Artificial Neural Networks

References

- Asef, M.R., and M. Farrokhrouz (2010), Methods and parameters for estimation of wave velocity in rock. **In:** *19th Int. Geophysical Congress and Exhibition, 23-26 November 2010, Turkey, Ankara*, DOI: 10.13140/2.1.5072.5440.
- Brocher, T.M. (2005), Empirical relations between elastic wavespeeds and density in the Earth's crust, *Bull. Seismol. Soc. Am.* **95**, 6, 2081-2092, DOI: 10.1785/0120050077.
- Castagna, J.P., M.L. Batzle, and R.L. Eastwood (1985), Relationships between compressional-wave and shear-wave velocities in elastic silicate rocks, *Geophysics* **50**, 4, 571-581, DOI: 10.1190/1.1441933.
- Farrokhrouz, M., and M.R. Asef (2010), Dominant parameters for estimation of shear wave velocity. **In:** *19th Int. Geophysical Congress and Exhibition, 23-26 November 2010, Turkey, Ankara*.
- Han, D.H. (1986), Effects of porosity and clay content on acoustic properties of sandstones and unconsolidated sediments, Ph.D. Thesis, Stanford University, Department of Geophysics, Stanford, USA.
- Liu, Y., Z. Chen, and K. Hu (2012), Shear velocity prediction and its rock mechanic implications. **In:** *Conf. GeoConvention 2012: Vision, 14-18 May 2012, Calgary, Canada*, 5 pp.
- Memari, A. (2013), Evaluation of surface subsidence in one of Iran's oil fields using INSAR technique, *Am. J. Oil Chem. Technol.* **1**, 4, 9-17.
- Rawlings, J.O., S.G. Pantula, and D.A. Dickey (1998), *Applied Regression Analysis: A Research Tool*, 2nd ed., Springer Texts in Statistics, Springer, New York.

Received 6 July 2014

Received in revised form 10 October 2014

Accepted 15 October 2014